PERFORMANCE ANALYSIS OF 8KBPS VOICE CODEC (G.729, G.711 ALAW, G.711 ULAW) FOR VOIP OVER WIRELESS LOCAL AREA NETWORK WITH RESPECTIVE SIGNAL-TO-NOISE RATIO

MOHAMMAD ZAKI BIN NORANI

THESIS SUBMITTED IN FULFILMENT OF THE DEGREE OF COMPUTER

SCIENCE (COMPUTER SYSTEM AND NETWORKING)

FACULTY OF COMPUTER SYSTEM AND SOFTWARE ENGINEERING

2013

ABSTRACT

In this research, 3 types of speech codec (G.729, G.711 aLaw and G.711 uLaw) in the same sampling rate of 8kbps are put to test in predefined network environment and given respective SNR 10dB, 20dB and 30dB to measure the performance base on R-factor, MOS, packet jitter and packet lost. Speech codec is used to convert the analog voice signals into digital signal. Each speech codec have its own speech quality, minimum bandwidth require etc. There are many manufacturers that have been producing various types of speech codecs in the market. The VoIP users are able to choose the desire codec that will be used or enable in the VoIP call based on the service and hardware that can support the speech codec. But, users will face some difficulty in choosing the best codec to use. All 3 mentioned speech codec will be test base on these criteria; VoIP session over optimum wireless network with 10dB, 20dB and 30dB SNR and VoIP session over wireless network that shared with other traffic with 10dB, 20dB and 30dB SNR. Six testbed will be carry out to complete all the criteria and all of the tests criteria will be carry out on real devices simulation. At the end, the performance measurement such as MOS, r-factor , packet lost and packet jitter will be observe to determine the best speech codec in each scenario. The final results of this research should be able to determine the best speech codec among the four codecs that have been selected and match the suitability with the environments.

Contents

Chapter 1

INTRODUCTION

1.0 Introduction

Voice over Internet Protocol (VoIP) is a technology that allows user to make voice calls using a broadband Internet connection instead of a regular (or analog) phone line. Some VoIP services may only allow user to call other people using the same service, but others may allow user to call anyone who has a telephone number including local, long distance, mobile, and international numbers. Also, while some VoIP services only work over user computer or a special VoIP phone, other services allow user to use a traditional phone connected to a VoIP adapter.

VoIP services convert user voice into a digital signal that travels over the Internet. If users are calling a regular phone number, the signal is converted to a regular telephone signal before it reaches the destination. VoIP can allow user to make a call directly from a computer, a special VoIP phone, or a traditional phone connected to a special adapter. In addition, wireless "hot spots" in locations such as airports, parks, and cafes allow user to connect to the Internet and may enable user to use VoIP service wirelessly. Usually the device and equipment that we need to use this service is a broadband (high speed Internet) connection is required. This can be over a cable modem, or high speed services such as DSL or a local area network. A computer, adaptor, or specialized phone is required. [1]

This experiment will analyze one of the parts of the VoIP which is speech codec. Speech codec are used to convert an analog voice signal to digitally encoded version. Codecs vary in the sound quality, the bandwidth required, the computational requirements, etc.

In VoIP, voice which is an analog signal is converted into digital signal, it is then encoded by using suitable VoIP codecs and your voice is compressed in the form of stream of binary data and transmitted over the internet. When encoded data arrives at the other end it is then decoded into digital signal and then analog signal. This whole process takes place within less than few milliseconds. [2]

1.2 Problem statement

With many codecs has been produced by various manufacturers, the selection of the codec to use to be a little difficult when user need to consider the appropriateness of a codec suitable with bandwidth of the network. Some VoIP client provides the codec that the user can select is manually. In this type of application, the codec selection is really important because it will give big effect to the VoIP session whether the session will provide a good or bad quality of voice.

Furthermore, certain speech codec have a same minimum bandwidth requirement. For example G.729, G.711 aLaw, and G.711 uLaw have 8kb minimum bandwith requirement. So, among this four codec, user cannot determine which one of the codec is the best for VoIP session.

Many studies have been made to analyze the performance of speech codec based on sampling method and research base on ideal network. However, not many researches have done on specific group of speech codec based on minimum bandwidth requirement and specific type of network.

User that use the VoIP service usually using the VoIP service while using the other internet application such as web browsing, file transfer etc. So, it will affect the performance of the connection of the VoIP session because it’s using the same network bandwidth.

1.3 Objective

This study was conducted to meet three objectives:

- To simulate (G.729, G.711 aLaw, and G.711 uLaw) speech codec of VoIP on predefine wireless mesh network.
- To simulate (G.729, G.711 aLaw, G.711 uLaw) speech codec on respective (10dB, 20dB, 30dB) signal-to noise ratio (SNR).
- To analysis (G.729, G.711 aLaw, G.711 uLaw) speech codec performance in term of MOS and R-factor.
- To suggest the best speech codec based on the MOS score and R-factor for the predefined wireless mesh network and respective SNR.

1.4 Scope

Due to the time and resources constraints, this research is limited to the following matter:

- The simulation is using a group of 8kbps speech codec. Four type of speech codec (G. G.729, G.711 aLaw, G.711 uLaw) to be analyzed.
- Using only SIP architecture environment.
- IEEE 802.11n wireless network connection will be used as a medium during the simulation. Two wireless access points (AP) will be used to establish the connection.
- Equipment: computer, WAP, SIP server, VoIP client and real test bit simulation
- The end of the simulation, the performance measurement base on MOS and R-factor can be generated to see the result of the experiment.

Thesis Organization

The research consists of five chapters:

Chapters 1 provide the overall overview of the thesis. Here, the problem statement will be introduced. Then based on the problem statement, the objective of the research is being defined. Lastly, chapter one also will explain about the research scope.

Chapter 2 introduces the hardware and software that will be used in this research project. It is mainly focuses on the performance of the bandwidth estimation tools. The literature review is organized in a way that readers can understand this.

Chapter 3 explains the methodology that will be used to carry out this research. The detail will be elaborated step by step process that is being used to complete the research.

Chapters 4 design the model or know as architecture that will be developed in order to perform the test. It then followed with the continuously design on data analysis.

Chapter 5 concludes all the chapters and the recommendations for future researchers explain most of the configurations of hardware and software involved in the research. Detail test result will be included in this chapter.

Chapter 2

2.0 LITERATURE REVIEW

2.1 Session initiate protocol (SIP)

SIP is a signaling protocol like to HTTP. It is a protocol that can setup and tear down any type of session. SIP call control uses Session Description Protocol (SDP) to describe the details of the call (i.e., audio, video, a shared application, codec type, size of packets, etc.). SIP uses a URI7 to identify a logical destination, not an IP address. The address could be a nickname, an e-mail address, or a telephone number. In addition to setting up a phone call, SIP can notify users of events, such as “I am online,” “a person entered the room,” or “e-mail has arrived.” SIP can also be used to send instant text messages.

PSTN-Like services	Create new services
Caller ID	Web/voice integration
PBX-like features	Programmable services
Call forwarding	Multi-destination routing
Call transfer	Presence
AIN-like features	Instant messaging
Free phone	Multimedia
Find me/follow me	Event notification
Conference calls	Caller and called party preferences

Table 1: defines some of the types of services that can be offered using SIP.

Using a client–server model, SIP defines logical entities that may be implemented separately or together in the same product. Clients send SIP requests, whereas servers accept SIP requests, execute the requested methods, and respond.

The SIP requirement defines six request methods [11]:

- **REGISTER** permits either the user or a third party to register communicates information with a SIP server.
- **INVITE** initiates the call signaling sequence.
- **ACK** and **CANCEL** sustenance session setup.
- **BYE** terminates a session.
- An **OPTION** queries a server about its abilities.

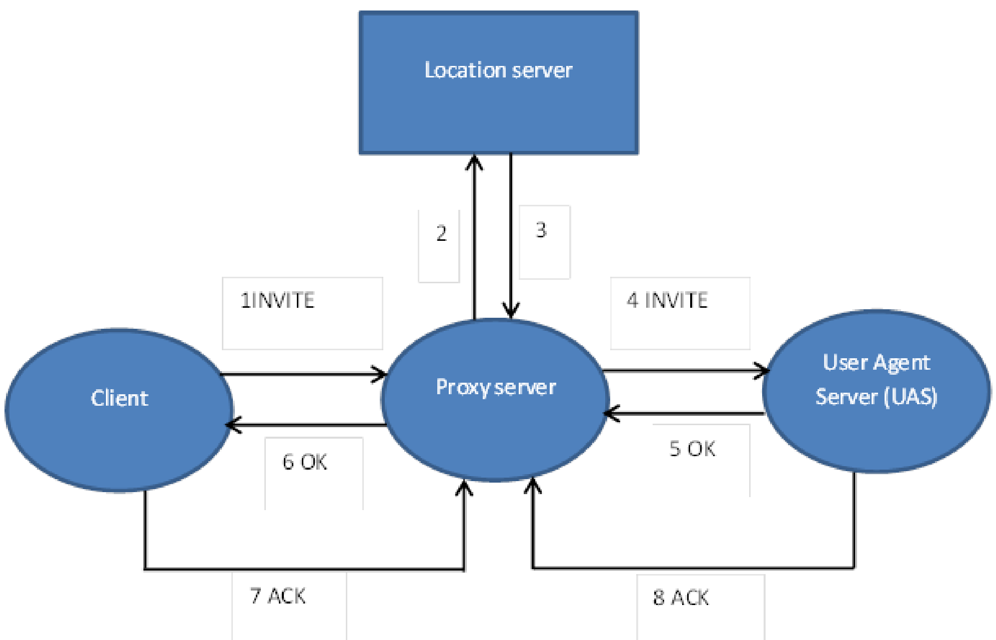

Figure 1: Basic example of a SIP operation

2.3 Speech codec

2.3.1 Overview

When we talk with people which nearby or there is at place which same with where we are, voice sent directly to ear listener through wave. To talk with people who are in other places, especially the different geographical areas such as using the telephone or other equipment, the sound transmitted through various mediums such as cable, microwave, and signal and so on. Therefore, we need communication equipment require human voice codecs to convert the analog form to digital form of the signal to be transmitted through the medium provided. Codecs are used to convert an analog voice signal to digitally encoded version. Codecs vary in the sound quality, the bandwidth required, the computational requirements, etc.

Many factors determine voice quality, including the choice of codec, echo control, packet loss, delay, delay variation (jitter), and the design of the network. Packet loss causes voice clipping and skips. Some codec algorithms can correct for some lost voice packets. Typically, only a single packet can be lost during a short period for the codec correction algorithms to be effective. If the end-to-end delay becomes too long, the conversation begins to sound like two parties talking on a Citizens Band radio. A buffer in the receiving device always compensates for jitter (delay variation). If the delay variation exceeds the size of the jitter buffer, there will be buffer overruns at the receiving end, with the same effect as packet loss anywhere else in the transmission path.

There are many codecs available for digitizing speech. The quality of a voice call through a codec is often measured by subjective testing under controlled conditions using a large number of listeners to determine an MOS. Several characteristics can be measured by varying the test conditions. Important characteristics include the effect of environmental noise, the effect of channel degradation (such as packet loss), and the effect of tandem encoding/decoding when interworking with other wireless and terrestrial transport networks. The latter characteristic is especially important since VoIP networks will have to interwork with switched circuit networks and wireless networks using different codecs for many years [11].

2.3.2 ITU-T G.729

G.729 is ITU-T codec standard that have two version which is A and B. G.729 also offered 8kbps low bit rate with reasonably toll-quality voice. Input frames are 10 milliseconds (10 ms) in duration and generated frames contain 80 bits. The input and output contain 16-bit pulse-code modulation (PCM) samples converted from or to 8-Kbps compressed data. Toll-quality is the service that this codec will provide is can give same quality with public switch network where the call is charged every minute of use [4]. Therefore, it ideally suited for the use of VoIP because VoIP broadband rates should be given serious consideration. G.729 is the proprietary codec; anyone that wants to use this codec on their client should get the license from the company that re-sells the G.729 license. However, there has some non-commercial experimental for this G.729 can be used.

G.729 fits into the general category of CELP (Code Excited Linear Prediction) speech coders [6]. These coders are all based on a model of the human vocal system. In that model, the throat and mouth are modeled as a linear filter, and voice is generated by a periodic vibration of air exciting this filter. In the frequency domain, this implies that speech looks somewhat like a smooth response (called the envelope), modulated by a set of discrete frequency components. CELP coders all vary in the manner in which the excitation is specified, and the way in which the coefficients of the filter are represented. All of them generally break speech up into units called frames, which can be anywhere from 1ms to 100ms in duration. For each frame of speech, a set of parameters for the model are generated and sent to the decoder. This implies that the frame time represents a lower bound on the system delay; the encoder must wait for at least a frames worth of speech before it can even begin the encode process. In G.729, each frame is 10ms, or 80 samples, in duration. This frame is further broken into two 5ms sub frames. The filter parameters are specified just once for each frame, but each sub frame has its own excitation specified. It is also important to note that speech can generally be classified into two types: voiced and unvoiced. Voiced sounds, such as b,d, and g, are generated from the throat, whereas unvoiced sounds, such as th, f, and sh, are generated from the mouth. The model works better for voiced sounds, but the excitation can be tailored for voiced or unvoiced so that it works in both cases.

2.3.3 G.711

An International Telecommunications Union (ITU-T) standard for audio (speech) compression and decompression that is used in digital transmission systems, and in particular, used for the coding of analog signals into digital signals.

G.711 is also known as Pulse Code Modulation (PCM). It is the ITU-T international standard for encoding telephone audio on a 64 kbps channel. PCM samples the signal 8000 times a second; each sample is represented by 8 bits for a total of 64 kbit/s. There are two versions of this standard codec. The u-law (pronounced as mew law) is generally used in North America and Japan digital communications. The A-law is used in European digital communications. The difference between the two standards is the method in which the analog signal is sampled. (See also PCM).

2.3.3.1 G.711 a-Law

A-law is used in Europe and the rest of the world. This type of G.711 has a smaller dynamic range than U-law. Dynamic range is basically the ratio between the quietest and loudest sound that can be represented in the signal. The downside of having a higher dynamic range is greater distortion of small signals. This simply means that a-law would sound better than u-law when the sound input is very soft.

In addition, this codec encoding thus takes a 13-bit signed linear audio sample as input and converts it to an 8 bit value as show in the table below:

Linear input code	**Compressed code**
s0000000wxyz`a	s000wxyz
s0000001wxyz`a	s001wxyz
s000001wxyz`ab	s010wxyz
s00001wxyz`abc	s011wxyz
s0001wxyz`abcd	s100wxyz
s001wxyz`abcde	s101wxyz
s01wxyz`abcdef	s110wxyz
s1wxyz`abcdefg	s111wxyz

Table 2 audio sample compressed form for a-Law

2.3.3.1 G.711 U-Law

U-Law is a companding algorithm, primarily used in the digital telecommunication systems of North America and Japan. Companding algorithms reduce the dynamic range of an audio signal. In analog systems, this can increase the signal-to-noise ratio (SNR) achieved during transmission, and in the digital domain, it can reduce the quantization error (hence increasing signal to quantization noise ratio). These SNR increases can be traded instead for reduced bandwidth for equivalent SNR. Different from a-Law, u-Law has higher dynamic range than a-Law. U-law encoding takes a 14-bit signed linear audio sample as input. Different from a-Law, u-Law increases the magnitude by 32 (binary 100000), and converts it to an 8 bit value as show in the table below:

Linear input code	Compressed code
s00000001wxyz`a	s000wxyz
s0000001wxyz`ab	s001wxyz
s000001wxyz`abc	s010wxyz
s00001wxyz`abcd	s011wxyz
s0001wxyz`abcde	s100wxyz
s001wxyz`abcdef	s101wxyz
s01wxyz`abcdefg	s110wxyz
s1wxyz`abcdefgh	s111wxyz

Table 3 audio sample compressed form for u-Law

2.2 Speech codec performance measurement

2.2.1 Mean opinion score (MOS) and R-Factor

MOS is the standard to rate the test of audio quality recommended by ITU-T. MOS is determined in one number, from 1 to 5, 1 being the worst and 5 the best. Originally, the test is involve the human that being held with some people sat in the quiet room and listening the audio and the rate it with the MOS. This is why the word “opinion” is used. Nowadays, the test no longer uses human, software to measure and quantify audio MOS was developed. This software is able to calculate the MOS for a test made on audio [1].

R-factor is the alternative way to calculate rate the quality of speech and audio. The function is same as MOS but the rate is different which is from 1 to 120. Because of the limit of rate is more than MOS, R-factor is become the most precise method to do the test of the quality of speech and audio. R-factor evaluates the user perception and all factors that will be effect the quality of the VoIP system. The rate is divided by two results which is network and user R-factor. Most of the user believes that the R-factor is more objective method than MOS [1].

Although r-factor can calculate better evaluation result of the testing, both of the method will be use and generated to get better judgment of the call quality.

Rating	ACR description – MOS	DCR description – DMOS
5	Excellent	Degradation not perceived
4	Good	Degradation perceived but not annoying
3	Fair	Degradation slightly annoying
2	Poor	Degradation annoying
1	Bad	Degradation very annoying

Table 4: MOS and R-factor value

2.3 Typical VoIP Problems

2.3.1 Packet jitter

Jitter is one of the QOS issue in the use of VoIP if the problem can no longer be controlled. Different with network delay, jitter does not happen because of the packet delay, but it is happen when the variation of packet delays occur. Jitter occurs when a packet should be delivered in a steady stream, but due to network problems packet sent not arrive right on time [13].

In VoIP conversation, VoIP endpoints try to control the jitter by increase the size of the packet buffer; jitter will causes delays in the VoIP conversation. The minimum variation of the packet is 150ms and if the variation becomes too high and exceeds the minimum variation, callers will notice the delay and will talking like walkie-talkie conversation. There are several steps you can take to deal with jitter on the network layer and application layer such as VoIP software, IP phones or specific VoIP adaptors. By definition, steps to reduce delays in the network to maintain the buffer is less than 150ms but the variation could not necessary removed. Although variation is not removed by the reduction in network delay, but it is still effective in reducing variation and this is not known by the caller. In addition, the setting for VoIP services as a priority and bandwidth shaping on the network can also reduce the variation in packet delay [1].

At the endpoint, it is crucial to optimize the jitter buffering. While better buffers reduce and eliminate the jitter, anything over 150ms remarkably affects the real quality of the conversation. Adaptive algorithms to manage buffer size depending on the present network circumstances are often quite working. Fiddling with the packet size or using a dissimilar codec (e.g. G.711) will often help control the jitter. While jitter is more affected by network delays than by the endpoints, some resource-struggling systems that are executed in concurrent environments, such as VoIP soft-phones, may present significant and random variations in packet delays. While developing VoIP endpoints or simulate call quality problems within the existing VoIP infrastructure, it is very important to consider about the cause of jitter. A network analyzing and monitoring tool with VoIP analysis can be use in localizing the source of the problem efficiently by produce the value of the jitter and packet loss of the VoIP session.

2.3.2 Packet loss

Packet loss can occur in all types of networks. Therefore, each network protocol designed to handle packet loss occurrences in their own ways. For example, the Transmission control protocol (TCP), which address the problem of packet loss in transmission of a packet with the request for packets that have been lost during the transmission. But in VoIP, VoIP call has no time to wait for the packet to arrive.

VoIP is very concerned about the problem of packet loss, even if only 1% of the packets have been dropped; it will affect the quality of VoIP calls [4]. Speech codecs will play an important role in dealing with the problem of packet loss. Most speech codecs can only assume less than 1% packet loss on a VoIP call this problem should be avoided during VoIP calls from the occurrence of audible errors. Ideally, there must be no packet loss in VoIP call.

There are several techniques to prevent or reduce the packet loss problem. One of the techniques that usually use is called Packet Loss Concealment (PLC) has been used in VoIP session to mask the effect of packet loss. There are some more techniques that may be used in different implementations.

In VoIP, some packets will be discarded for several reasons, including network congestion, line errors, and late arrival. Look at the exact value of packet-loss graphs permit the network administrators to select a PLC technique that best counterparts the characteristics of a certain environment, this method will help them to manage the problem of packet loss effectively.

2.3.3 Signal-to-noise ratio (SNR)

Signal-to-noise ratio (SNR) is one of the factor that VoIP developer should consider when develop VoIP over wireless link. The signal to noise ratio is also referred to as SNR. In other word, it is the ratio between the maximum signal strength that a wireless connection can reach and the noise present in the connection. Here “noise” refers to the stray frequencies that obstruct with the transmission of data in a wireless network.

The signal-to-noise ratio, the bandwidth, and the channel capacity of a communication channel are related by the Shannon–Hartley theorem. Signal-to-noise ratio is sometimes used indirectly to refer to the ratio of useful info to false or unrelated data in a conversation or exchange. The Shannon–Hartley theorem states the maximum rate at which information can be transmitted over a communications channel of a specified bandwidth in the occurrence of noise. Therefore, the SNR of a network have to be as high as possible. The greater the value of SNR, the better the signal strength and the quality of transmission will get.

In VoIP over wireless LAN environment the quality of the call session is depending on how good the voice transmission will be conducted. Based on the Shannon-Hartley theorem, when the SNR interrupt or present on the wireless link that will be used for VoIP session, the call quality and voice data transmission will affected. The value of the SNR can also influence the quality of call and affect the MOS and R-Factor reading in every call because.

2.4 Related research

2.4.1 Capacity of an IEEE 802.11b Wireless LAN supporting VoIP

This research evaluate the capacity of an IEEE 802.11b network carrying voice calls in a wide range of scenarios, including varying delay constraints, channel conditions and voice call quality requirements. It uses G.711 and G.729 voice encoding schemes and a range of voice packet sizes. Firstly, researcher present an analytical upper bound and, using simulation, show it to be tight in scenarios where channel quality is good and delay constraints are weak or absent. Then use the simulation to show that capacity is highly sensitive to the delay budget allocated to packetization and wireless network delays.

This research also shows how channel conditions and voice quality requirements affect the capacity. Selecting the optimum amount of voice data per packet is shown to be a trade-off between throughput and delay constraints: by selecting the packet size appropriately given the delay budget and channel conditions, the capacity can be maximized. Unless a very high voice quality requirement precludes its use, G.729 is shown to allow a capacity greater than or equal to that when G.711 is used, for a given quality requirement [2]. The paper is only evaluated an upper bound on the capacity of IEEE 802.11b network carrying voice calls, and found it to be tight in scenarios where channel quality is good and delay constraints are weak or absent [2]. Furthermore, this paper only uses two types of codec which is G.711 and G.729 that have different bit rate. The test is manipulating the network environment such as delay, packet size and channel condition to test the performance and quality of speech in the conversation [2].

2.4.2 Implementing VoIP: A Voice Transmission Performance Progress Report

This paper is aiming to introduce voice over IP networks and services in ways that satisfy the voice quality expectations of one of the VoIP service provider company customers, they have been conducting laboratory studies of how VoIP transmission affects voice quality while also carefully monitoring and managing several field implementations of VoIP. This article summarizes much of what they have learned in this work, and they hope it provides a useful progress report on the industry's evolution to VoIP.

They review their data on the voice quality effects of packet loss, delay, speech coders, packet loss concealment algorithms, and the compression option of suppressing transmission during silence. Because the familiar problem of echo has emerged repeatedly in the VoIP environment, they review this issue in some detail. Packet loss and delay variation measurements made on private VoIP networks are reviewed, and the data here are encouraging [3]. The test recommended for VoIP session, the codec that practical to use to maintain the performance is G.711, G.726, G.728, or G.729E [3]. This research also use the single VoIP session in the network environment without any other traffic from any network application that use the same network environment. The paper focus on the type of service that has been provides by the service provider to their customer. It suggest a few method to get better service to be provide to customer such as compression can come with significant quality penalties, especially where multiple coding are likely end-to-end and/or where high noise levels (or music on hold) might he a common operating condition.

2.4.3 VoIP Basics: Codec Latency vs. Bandwidth Optimization

This paper is about the codec latency and bandwidth optimization. The researcher has indicated that the codec has low bandwidth is very efficient. This is proved by G.729 will compress 10 milliseconds of audio to 10 bytes and G.723.1 encode 30ms frames to 24 or 20 bytes [9]. However, since we send compressed audio frames as payload in RTP packets which are in turn sent over UDP, The researcher need to consider the overhead for IP, UDP, and RTP headers. The overhead is 40 bytes per packet. This is significant when compared with the size of a compressed audio frame if we are not on a local area network and the bandwidth is limited. The table below shows the overhead for several low-bandwidth codecs. The researcher did the calculation for one frame per packet for G.723.1 and GSM and for 3 frames per packet for G.729 since this codec works with frame size of only 10 milliseconds.

Codec	Nominal bitrate [kbit/s]	Frame length [ms]	Frame size [bytes]	Packet overhead	Actual bitrate [kbit/s]
G.723.1	6.4	30	24	167%	17
G.723.1	5.3	30	20	200%	16
G.729	8	10 *3	10 *3	133%	18.6
GSM 06.10	13	20	33	121%	29.2

Table 5: calculation for one frame per packet for G.723.1 and GSM and for 3 frames per packet for G.729

When calculating the latency, you need to consider the time it takes to send a packet from one end to another (your mileage may vary, try to use "traceroute" to get a clue) and the size of the jitter buffer of the receiving end (which can be 50-60 milliseconds worth of audio). Considering all this, I would say the reasonable maximum is to send 60 milliseconds of audio in one packet. This will result in the following bitrates:

Codec	Nominal bitrate [kbit/s]	Frames in 60 ms of audio	Actual bitrate [kbit/s]
G.723.1	6.4	2	11.7
G.723.1	5.3	2	10.6
G.729	8	6	13.3
GSM 06.10	13	3	18.5

Table 6: result of 60 milliseconds of audio in one packet

In addition to latency, there are two more things we should consider when increasing the number of audio frames per RTP packet:

- If a packet with a larger number of frames gets lost, the loss is more noticeable to the user.
- With greater end-to-end delay, possible echo become more noticeable.

2.4.4 Performance Analysis of Different Codecs in VoIP Using SIP

Converged IP networks look for to incorporate voice, data, and video on the same infrastructure. Nevertheless, the integration of all kinds of traffic onto a single IP network has some benefits as well as weaknesses. While decreasing cost and growing mobility and functionality, VoIP may lead to consistency concerns, degraded voice quality, incompatibility, and end-user complaints due to moving network characteristics. The main purpose of VoIP, various codecs used in VoIP and packet loss, Jitter, delay are analyzed and discussed.

The comparison between three different codecs which is G.711, G.723 and G.729 has been analyzed by implementing peer-to-peer VoIP network using SIP server and caller. Thus we have described the various codecs in VoIP implementation and analyzed three commonly used codecs using peer-to-peer network scenario. These are common narrow band codecs. It can be analyzed from the results that G.711 is an ideal solution for PSTN networks with PCM scheme. G.723 is used for voice and video conferencing however provides lower voice quality. Music or tones such as DTMF cannot be transmitted reliably with G.723 codec. G.729 is mostly used in VoIP applications for its low bandwidth requirement.

Codec	*Data Rate*	*MOS Score*
G 711	64	4.3
G 726	32	4.0
G 726	63	3.8
G 728	16	3.9
G 729	8	4.0
GSM	13	3.7

Table 7: MOS for this analysis

2.4.5 An E-Model Implementation for Speech Quality Evaluation in VoIP Systems

The most common method is very effective for measuring the quality of voice and sound is to use MOS. however, the combination of MOS and R-factor enhance voice quality test results will be made. This article presents a voice quality measurement tool based on the ITU-T E Model. Firstly, the

ITU-T and ETSI specifications of E-Model are briefly reviewed and some errors found in these documents are pointed [12]. After, a measurement tool based on the corrections is described. E-Model can each impairment factor which affects a voice call can be computed separately, even so this does not imply that such factors are uncorrelated, but only that their contributions to the estimated impairments are separable. An expressive amount of delay and lost packets have to be present in the call, at alternating burst and gap conditions, otherwise we always will have excellent MOS scores [12].

In order to validate the measurement tool operation, we have to generate VoIP calls under known QoS environments and evaluate its voice quality using the tool. An expressive amount of delay and lost packets have to be present in the call, at alternating burst and gap conditions, otherwise we always will have excellent MOS scores and the measurement tool would not be completely tested. Thus we used the scenario shown on Figure 2 to generate some calls that could be evaluated by the measurement tool [12].

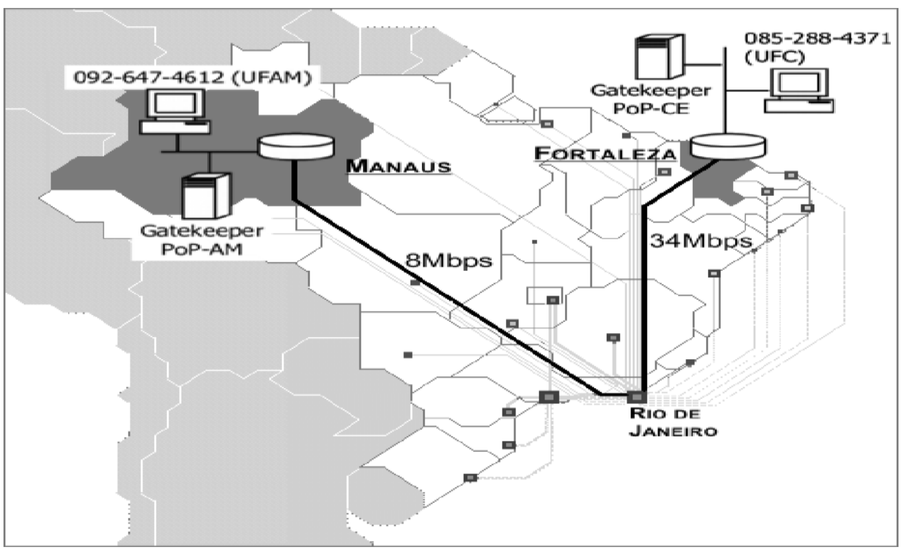

Figure 2 the diagram of the scenario that the researcher created

2.4.6 Best VoIP codecs selection for VoIP conversation over wireless carriers' network

This research is about to determine the best performance of VoIP using different codec. The impact of each element of a VoIP call will be analyzed. The rates will be generated from these VoIP elements will help in determining the best codec between five codec selected (G.711, G.722, G.726, GSM and SPEEX). Below is a two environment that has been use for this research:

1. VoIP over wireless LAN
    - All five codec has been analyze in this environment to be able to determine the best codec if the VoIP session is using the wireless LAN network.

2. VoIP over wireless WAN
    - The codec that has been selected is tested using the wireless WAN as a communication medium. The two different of the WAN ISP has been selected and used for this test.

In the analysis phase, the researcher measures and compares the VoIP performance using different codecs selection. The generating of the packet loss, packet jitter and MOS will help the researcher to determine the best codec that suitable to use in each environment. Five reading has been made to get the most accurate value. The specific tool is used to get the value of the packet loss, packet jitter and MOS.

2.4.7 Comparison between related researches

Research title	Codec used	Method	Result/Analysis
Capacity of an IEEE 802.11b Wireless LAN supporting VoIP	- G.711 - G.729	- Research evaluate the capacity of an IEEE 802.11b network carrying voice calls in a wide range of scenarios, including varying delay constraints, channel conditions and voice call quality requirements.	- Evaluated an upper bound on the capacity of an IEEE 802.11b network carrying voice calls. - the use of G.729 has been shown to allow - Greater capacity than the use of